

Multi-plasmon absorption in graphene

Marinko Jablan^{1,2,*} and Darrick E. Chang¹

¹*ICFO-Institut de Ciències Fotoniques, Mediterranean Technology Park, 08860 Castelldefels (Barcelona), Spain*

²*Department of Physics, University of Zagreb, 10000 Zagreb, Croatia*

(Dated: January 22, 2015)

We show that graphene possesses a strong nonlinear optical response in the form of multi-plasmon absorption, with exciting implications in classical and quantum nonlinear optics. Specifically, we predict that graphene nano-ribbons can be used as saturable absorbers with low saturation intensity in the far-infrared and terahertz spectrum. Moreover, we predict that two-plasmon absorption and extreme localization of plasmon fields in graphene nano-disks can lead to a plasmon blockade effect, in which a single quantized plasmon strongly suppresses the possibility of exciting a second plasmon.

PACS numbers: 73.20.Mf, 78.67.Wj, 79.20.Ws, 42.50.Hz

The field of nonlinear optics ranges from fundamental questions concerning light-matter interactions to exciting technological applications [1]. However, usually very large field intensities are required to observe nonlinear effects. One is thus always looking for systems that will exhibit nonlinear phenomena at lower powers, with the ultimate limit being strong interactions between just two quanta of light [2]. One possibility to increase nonlinear effects is to use the strong localization and enhancement of electromagnetic fields in the form of surface plasmon excitations [3]. In that regards, we note that graphene [4] has been demonstrated to support extremely localized plasmons [5–12]. While optical nonlinearities in graphene have been studied by several authors [13–24], here we predict a novel nonlinear effect in the form of multi-plasmon absorption. We also show how this effect leads to saturable absorption in graphene nano-ribbons at low input powers in the far-infrared and terahertz spectrum. Moreover, we predict that the extreme localization of plasmon fields in graphene nano-disks leads to such a strong two-plasmon absorption that it becomes nearly impossible to excite a second plasmon quanta in the system. This plasmon blockade effect would cause the nano-disk to behave essentially like a quantum two-level system, which is observable in its resonance fluorescence spectrum.

Graphene is a two-dimensional hexagonal lattice of carbon atoms [4]. The low-energy band structure of graphene is described by Dirac cones with the electron dispersion $E_{n\mathbf{k}} = n\hbar v_F |\mathbf{k}|$, where $v_F = 10^6$ m/s, and $n = \pm 1$ stands for the conduction (valence) band [25]. In its intrinsic form graphene is a zero-gap semiconductor; however, it can also be easily doped with free carriers and as such it supports plasmon modes [5–7]. At low frequencies, one can get a rather accurate description of these modes by using a simple Drude conductivity,

$$\sigma(\omega) = \frac{e^2 E_F}{\pi \hbar^2} \frac{i}{\omega + i\gamma}, \quad (1)$$

where E_F is the Fermi energy of graphene and γ is a phenomenological damping rate that takes into account various decay channels like impurity or phonon scatter-

ing [7]. The resulting plasmon dispersion is given by

$$q = \frac{2\pi\epsilon_0\bar{\epsilon}_r\hbar^2}{e^2 E_F} \omega^2, \quad (2)$$

and we assume the average dielectric permittivity $\bar{\epsilon}_r \approx 2.5$, which roughly corresponds to the case of graphene on SiO₂ substrate and air on top. This simple Drude model breaks down at large frequencies when plasmons become strongly damped by electron-hole excitations, which can be described by the Random Phase Approximation (RPA) [5–7]. However, at low energies the Pauli principle blocks this decay channel and the plasmon is a long-lived excitation [5–12]. The resulting plasmon wavelength $\lambda_p = 2\pi/q_p$ is about 100 times smaller than the free space wavelength $\lambda = 2\pi c/\omega$, leading to the extreme localization of electromagnetic fields [7].

An intuitive explanation of the strong nonlinearities associated with plasmons emerges by considering the typical doping levels in graphene. For an electron density of $n = 10^{12}$ cm⁻², the distance between two electrons is $r_e = 1/\sqrt{n\pi} = 5.6$ nm, and so to observe some kind of nonlinear phenomena we need to compete with an intrinsic electric field of the order

$$E_e = \frac{e}{4\pi\bar{\epsilon}_r\epsilon_0 r_e^2} \approx 2 \cdot 10^7 \text{ V/m}. \quad (3)$$

This is significantly smaller than the characteristic field amplitude associated with nonlinear effects in atoms [1], given by the field between an electron and proton at a distance of a Bohr radius $a_B = 0.5$ Å. The electric field in that case is $E_{at} = 5 \cdot 10^{11}$ V/m, about 4 orders of magnitude larger than the field E_e in graphene!

To see what happens to plasmons at such a field strength E_e , let us imagine a general case of plasma oscillations at frequency ω and wavevector q , which is accompanied by an electric field $E(\mathbf{r}, t) = \frac{1}{2} E_p e^{i\mathbf{q}\cdot\mathbf{r} - i\omega t} + c.c.$ in the plane of graphene. When the plasmon field is small, it will induce a surface charge density $\rho_p = \frac{q\sigma(\omega)}{\omega} E_p$ that in turn creates an electric field $E_p = \frac{-i\rho_p}{2\bar{\epsilon}_r\epsilon_0}$, thus driving self-sustained charge density oscillations. However, at

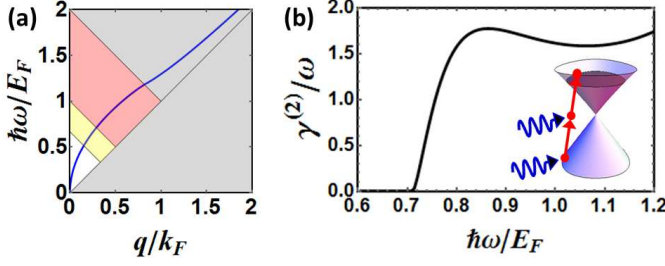


FIG. 1: (a) Plasmon dispersion relation. Gray area denotes the regime where a single plasmon can excite an e-h pair. Red area denotes the regime where this process is forbidden but plasmons can decay via two-plasmon absorption. Similarly, the yellow area denotes the regime of three-plasmon absorption. (b) Two-plasmon damping rate $\gamma^{(2)}/\omega = F^{(2)}(\omega)$ for $E_p = E_e$, indicating that plasmons cannot oscillate when the plasmon field is equal to the intrinsic electric field. Inset shows the band structure of graphene and a two-plasmon absorption process.

the field strength $E_p = E_e$, the induced charge density $|\rho_p| = en/2$ will be comparable to the initial charge density en , and this simple picture of plasmons breaks down. We will in fact see that at this field E_e there is a strong plasmon damping via the multi-plasmon absorption.

To understand how this comes about, let us first look at the linear (single-plasmon) absorption. Assuming that the graphene plane is perpendicular to the z axis, the plasmon field can be described by the scalar potential $\varphi(\mathbf{r}, z, t) = \frac{1}{2}\varphi_p e^{i\mathbf{q}\cdot\mathbf{r} - q|z|} e^{-i\omega t} + c.c.$, where $\mathbf{r} = (x, y)$. The electric field is then given by $\mathbf{E} = -\nabla\varphi$, with the amplitude $|E_p| = q|\varphi_p|$, for both in-plane and out-of-plane components. The interaction of the plasmon field with the electrons can be described by the Hamiltonian $H_p = -\frac{1}{2}e\varphi_p \sum_j e^{i\mathbf{q}\cdot\mathbf{r}_j} e^{-i\omega t} + c.c.$, where $\hat{\mathbf{r}}_j$ is the position operator acting on the j^{th} electron. To calculate linear absorption we can write the dissipated power as $P = \sum_n \hbar\omega \frac{dw_{n0}^{(1)}}{dt}$, where

$$\frac{dw_{n0}^{(1)}}{dt} = \frac{2\pi}{\hbar} |\langle n|H_p|0\rangle|^2 \delta(E_n - E_0 - \hbar\omega) \quad (4)$$

is the Fermi golden rule for the probability that the system will absorb one plasmon quantum of energy $\hbar\omega$. Here $|n\rangle$ is the many-body excited state of momentum $\hbar\mathbf{q}$ and energy $\hbar\omega$, with respect to the ground state $|0\rangle$, and we assume that the system is at zero temperature [26]. To quantify absorption we can look at the dissipation rate,

$$\gamma = \frac{P}{W}. \quad (5)$$

Here $W = \frac{1}{2} \int d^2r \rho(\mathbf{r})\varphi(\mathbf{r})$ is the total electrostatic energy stored in the system, which is given by $W = \frac{1}{2}\epsilon_r\epsilon_0|E_p|^2 A/q$, where A is the surface area of the graphene flake. Finally, the relevant figure of merit is the plasmon quality factor $Q = \omega/\gamma$.

From Eq. (4) we see that the absorption process consists of a sum over all events where the plasmon can excite a single e-h pair. Conservation of energy and momentum require that $\hbar\omega = E_{n_2\mathbf{k}+\mathbf{q}} - E_{n_1\mathbf{k}}$; however, the Pauli principle allows this process only above the threshold condition of $\hbar\omega > 2E_F - \hbar v_F q$ (see the gray area in Figure 1 (a)). We can calculate the linear absorption by using the basis of Dirac electron wavefunctions in graphene [26]. As an example, at energy $\hbar\omega = 1.5 E_F$, the dissipation rate $\gamma^{(1)}/\omega = 1.1$ is so high ($Q = 0.9$) that the plasmon is not a well-defined excitation. Below this threshold the Pauli principle blocks the absorption process and the plasmon is a long-lived quasi-particle. However, if we increase the plasmon field, higher-order (nonlinear) absorption must be accounted for as well.

We note that this simple calculation of the linear absorption gives the same result as from the RPA [26]. Encouraged by this fact, we proceed to calculate the nonlinear, two-plasmon absorption by writing $P = \sum_n 2\hbar\omega \frac{dw_{n0}^{(2)}}{dt}$ and using the Fermi golden rule for the probability that the system absorbs two plasmon quanta:

$$\frac{dw_{n0}^{(2)}}{dt} = \frac{2\pi}{\hbar} \left| \sum_m \frac{\langle n|H_p|m\rangle \langle m|H_p|0\rangle}{E_m - E_0 - \hbar\omega} \right|^2 \delta(E_n - E_0 - 2\hbar\omega). \quad (6)$$

Alternatively one could perform third order expansion of the single particle density matrix including the screening fields consistently in every order of the expansion [27]. Such a calculation yields additional terms that contribute to the dissipation, arising from higher-order screening correlations. However, at high fields, when the nonlinear absorption is large, the screening process will be less effective and the simple calculation (6) should give a good estimate of the absorption. By evaluating expression (6) we get the two-plasmon absorption rate

$$\frac{\gamma^{(2)}}{\omega} = F^{(2)}(\omega) \left| \frac{E_p}{E_e} \right|^2, \quad (7)$$

where $F^{(2)}$ is a dimensionless function of plasmon frequency, which is given in the Supplemental Material [26] and shown in Figure 1 (b). It is straight forward to show that the Pauli principle allows the two-plasmon absorption only above the threshold $\hbar\omega > E_F - \hbar v_F q$ (see the red area in Figure 1 (a)). Then if we look at the energy $\hbar\omega = E_F$, single plasmon can not excite an e-h pair but it can decay via two-plasmon absorption. At this energy $F^{(2)} = 1.6$ and we see that intrinsic field E_e sets a natural scale for the nonlinear effects. Specifically at the field $E_p = E_e$ the two-plasmon absorption rate $\gamma^{(2)}/\omega = 1.6$ is so large that plasmon is not a well defined excitation. In fact at this intrinsic field we expect that the whole perturbation theory should fall apart. Indeed at the same energy $\hbar\omega = E_F$ and field strength $E_p = E_e$ we obtain the three-plasmon absorption rate $\gamma^{(3)}/\omega = 0.7$ [26]; which is about 40% of the two-plasmon

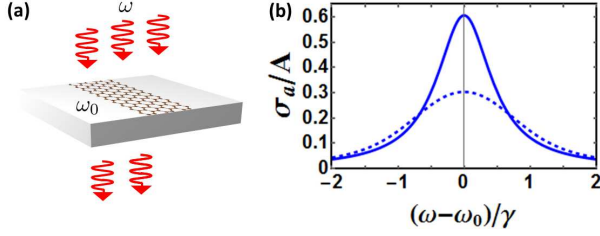


FIG. 2: (a) Transmission spectroscopy of a graphene nano-ribbon. (b) Absorption cross section normalized to the surface area of the ribbon. Solid line stands for the low pump $I \ll I_s$, and dashed line for the high pump intensity $I_s = 6 \text{ kW/cm}^2$. The saturation of absorption is caused by two-plasmon absorption inside of the ribbon. The ribbon width is $D = 25 \text{ nm}$, doping $n = 10^{12} \text{ cm}^{-2}$, and resonance frequency $\omega_0 = 2\pi \times 28 \text{ THz}$.

absorption rate, signalling the breakdown of the perturbation theory. Moreover, at the same field strength but lower energy $\hbar\omega = 0.6 E_F$ (in the yellow area in Figure 1 (a), where two-plasmon absorption is forbidden by the Pauli principle) we obtain the three-plasmon absorption rate $\gamma^{(3)}/\omega = 6.5$ again showing that plasmon can not oscillate at the intrinsic field E_e .

These effects could be easily observed in experiments involving extended graphene as the broadening of the plasmon linewidth with increasing field amplitude. However, we will now show that much more exciting behaviour can be seen by using graphene nano-structures to provide a resonance and field enhancement for the plasmon modes. The response of graphene nano-structure can be described by the polarizability [28, 29]:

$$\alpha(\omega) = D^3 \frac{G}{\frac{K}{\varepsilon_r} - \frac{4\pi\varepsilon_0 i \omega D}{\sigma(\omega)}}. \quad (8)$$

Particularly, in the case of a nano-ribbon, $K = 16$ and $G = L/D$, where D is the width and $L \gg D$ is the length of the ribbon [29]. To describe the plasmon resonance we can use the simple Drude model of the surface conductivity $\sigma(\omega)$ given by Eq. (1). However, now we must include the total damping rate $\gamma_t = \gamma + \gamma_{nl}$, which contains both the linear term (γ) like impurity or phonon scattering, and the nonlinear term (γ_{nl}) like two- or three-plasmon absorption. Specifically, to produce a resonance at frequency ω_0 , we require a ribbon of width: $D = \frac{K}{4\pi^2} \lambda_p \approx \lambda_p/2$, where $\lambda_p = 2\pi/q_p$ is the plasmon wavelength in the extended graphene given by Eq. (2). We are primarily interested in the regime $\hbar\omega_0 = E_F$ where a single plasmon can not excite an e-h pair, but two-plasmon absorption is allowed. For a typical doping $n = 10^{12} \text{ cm}^{-2}$ ($E_F = 0.12 \text{ eV}$), the corresponding free-space wavelength is $\lambda_0 = 2\pi c/\omega_0 = 10.6 \text{ }\mu\text{m}$, plasmon wavelength $\lambda_p = 62 \text{ nm}$, and a ribbon width $D = 25 \text{ nm}$.

The absorption cross-section is given by $\sigma_a(\omega) = 4\pi\frac{\omega}{c}\text{Im}\alpha(\omega) = \frac{3}{4\pi}\lambda_0^2 \frac{\gamma_r \gamma_t/2}{(\omega_0 - \omega)^2 + \gamma_t^2/4}$, where we have in-

troduced the radiative decay rate of the ribbon $\gamma_r = \frac{16\pi^3}{3} \frac{\varepsilon_r G}{K} \frac{D^3}{\lambda_0^3} \omega_0$ [29, 30]. To estimate impurity or phonon scattering we can use measurements of direct current mobility $\gamma = \frac{e v_F}{\mu \hbar \sqrt{\pi n}}$ [7]. For typical graphene mobilities of $\mu = 10^4 \text{ cm}^2/\text{Vs}$ [4], we have $\gamma = 8.6 \text{ THz}$, while $\gamma_r = 6.1 \text{ GHz}$, even for a ribbon of length $L = 100 D$. Since $\gamma_r \ll \gamma$, graphene nano-structures primarily act as absorbers, while the absorbed power can easily be detected via the reduction of light transmitted across the ribbon [9, 10] (see Figure 2 (a)).

Dissipated power is $P = I\sigma_a = \gamma_t W$, where W is the total energy, and $I = \frac{1}{2}\varepsilon_0 c |E|^2$ is the incident light intensity. We can then estimate the plasmon field inside of the ribbon by using the result for the extended graphene $W = \frac{1}{2}\varepsilon_r \varepsilon_0 |E_p|^2 A/q_p$, where $A = LD$ is the ribbon area.

We obtain: $|E_p|^2 = \frac{1}{2} \frac{\omega_0^2}{(\omega_0 - \omega)^2 + \gamma_t^2/4} |E|^2$, where the total damping rate $\gamma_t = \gamma + \gamma_{nl}$ itself depends on the plasmon field through the nonlinear damping term $\gamma_{nl} \propto |E_p|^2$ given by Eq. (7). Then by increasing the intensity, there is an increase in the total damping rate, and a decrease of the absorption cross section (see Figure 2 (b)). In particular, at intensity $I_s = 6 \text{ kW/cm}^2$, we obtain $\gamma_{nl} = \gamma$ on resonance, which reduces the absorption cross section by a factor of 2. This corresponds to an input power of only 2 mW for a laser focused to a diffraction-limited spot size $(\lambda/2)^2$, which would induce negligible heating of the graphene flake [26]. On the other hand, extended graphene can also be used as a saturable absorber, due to Pauli blocking, but the required intensities are 4 orders of magnitude higher, $I_s = 0.3 \text{ GW/cm}^2$ [16]. Moreover, by lowering the doping we can further reduce the saturation intensity of the ribbon, while the resonance frequency would fall into the terahertz spectrum.

Especially interesting is the case of a nano-disk where we can localize the entire field to an extremely small volume $\approx \lambda_p^3$. The polarizability of a disk can also be described by the expression (8), if we now take D as the disk diameter, $K = 12.5$ and $G = 0.65$ [28]. To produce a resonance at energy $\hbar\omega_0 = E_F$ and doping $n = 10^{12} \text{ cm}^{-2}$, requires a disk diameter $D = 20 \text{ nm}$. This yields a radiative decay rate of $\gamma_r = 24 \text{ MHz}$, while the other parameters are the same as in the case of a ribbon (above).

We can now estimate the electric field amplitude associated with a single quantized plasmon by writing $W = \hbar\omega_0$, and using the result for the extended graphene $W = \frac{1}{2}\varepsilon_r \varepsilon_0 |E_p|^2 A/q_p$, where $A = \pi D^2/4$ is the disk area. This gives a remarkable result for the field amplitude,

$$|E_p^Q| = \sqrt{\frac{2^8 \pi^4}{K^2} \frac{\hbar\omega_0}{\varepsilon_0 \varepsilon_r \lambda_p^3}} \approx 2 \cdot 10^7 \text{ V/m}. \quad (9)$$

In other words, the field of a single quantized plasmon is of the same order of magnitude as the intrinsic field E_e ! Then, for two plasmons the damping rate due to two-plasmon absorption would be so high ($\gamma^{(2)}/\omega_0 = 5.6$)

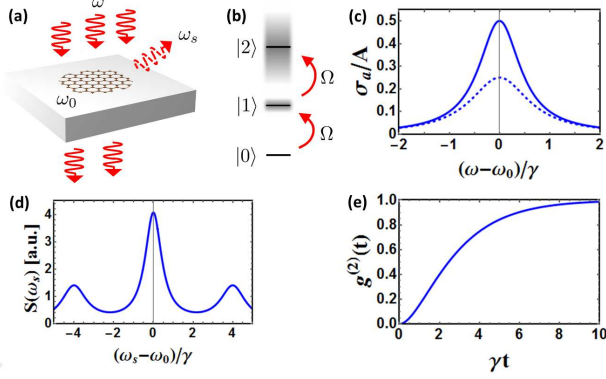


FIG. 3: (a) Resonance fluorescence of a graphene nano-disk. (b) Quantized plasmon energy levels in a disk. Two-plasmon absorption induces a huge linewidth of the doubly-excited state $|2\rangle$, which makes it extremely difficult to populate this state and effectively turns the disk into a two-level system. (c) Absorption cross-section normalized to the disk surface. The solid line stands for the low pump field ($\Omega \ll \gamma$), and the dashed line for the strong pump ($\Omega = \gamma/\sqrt{2}$). (d) Scattered spectrum on resonance ($\omega = \omega_0$) in the strong pump regime ($\Omega = 4\gamma$), clearly showing Rabi sidebands at $\omega_0 \pm \Omega$. (e) The second-order correlation function $g^{(2)}(t)$ of the scattered field for weak pump intensities, showing an anti-bunching effect $g^{(2)}(0) \ll 1$ characteristic of a two-level system.

that the resonance peak would completely disappear, leading to a plasmon blockade effect.

To quantify this effect we adopt a density matrix approach, $\frac{d\rho}{dt} = \frac{i}{\hbar} [\rho, H] + \mathcal{L}[\rho]$. Here the system Hamiltonian is given by $H = \hbar\omega_0 N - \frac{\hbar\Omega}{2}(ae^{i\omega t} + a^\dagger e^{-i\omega t})$, where $N = a^\dagger a$ is the plasmon number operator, and $\Omega = \sqrt{\frac{3}{2\pi} \frac{\lambda_0^2 I}{\hbar\omega_0} \gamma_r}$ is the Rabi frequency describing the interaction with the incident field of intensity I . The Liouvillian $\mathcal{L} = \mathcal{L}^{(1)} + \mathcal{L}^{(2)}$ consists of two terms that characterize the linear and nonlinear dissipation, respectively, $\mathcal{L}^{(N)}[\rho] = \frac{\gamma}{2}(2a^N \rho a^{N\dagger} - a^{N\dagger} a^N \rho - \rho a^{N\dagger} a^N)$ [31], where $\gamma^{(1)} = \gamma$ and $\gamma/\omega_0 = 0.05$, so that $\gamma^{(1)} \ll \gamma^{(2)}$. Under conditions of weak external driving, the dynamics of the (infinite-dimensional) density matrix can effectively be truncated to a few-excitation manifold [31]. Specifically at $\Omega \ll \gamma^{(2)}$ the steady-state population of the excited state $|2\rangle$ (containing two plasmons) is extremely weakly populated, $\rho_{22} = \frac{2\rho_{11}}{1+(2\gamma^{(2)}/\Omega)^2}$, and disk effectively behaves as a two-level system.

The absorption cross-section is then given by $\sigma_a = \frac{\hbar\omega_0\Omega}{I} \text{Im}\rho_{10} = \frac{3}{4\pi} \lambda_0^2 \frac{\gamma_r \gamma/2}{(\omega_0 - \omega)^2 + \gamma^2/4 + \Omega^2/2}$, which saturates at intensity $I_s = 50 \text{ kW/cm}^2$ ($\Omega = \gamma/\sqrt{2}$). Like in the case of the ribbons one finds that the radiative damping rate is negligible, $\gamma_r \ll \gamma$, and the disk primarily acts as an absorber. However, the weakly scattered light will now show interesting spectral properties, since it is well-known that a two-level system behaves as a strong frequency mixer [31]. To substantiate this, let us look

at the power spectrum of the scattered light, $S(\omega_s) = \frac{1}{\pi} \text{Re} \int_0^\infty \langle a^\dagger(0)a(t) \rangle e^{i\omega_s t} dt$. By using the quantum regression theorem [31] we obtain the steady-state two-time correlation function on resonance, $\langle a^\dagger(0)a(t) \rangle = \frac{1}{4} (e^{-\gamma t/2} + \frac{1}{2} e^{-3\gamma t/4} e^{-i\Omega t} + \frac{1}{2} e^{-3\gamma t/4} e^{i\Omega t}) e^{-i\omega t}$, where we have assumed the strong-pump regime $\Omega \gg \gamma/4$, but also $\Omega \ll \gamma^{(2)}$ so that the disk still behaves as a two-level system. The scattered spectrum consists of three peaks; one at the laser frequency ω_0 and two at the Rabi sidebands $\omega_0 \pm \Omega$. In Figure 3 (d) we plot the case of $\Omega = 4\gamma$, while $\Omega/\gamma^{(2)} = 0.04$ so that the disk still behaves as a two-level system up to an excellent approximation. A peculiar feature of this result is that the system produces frequency mixing (characteristic of dispersive nonlinearities) starting only from the two-plasmon absorptive nonlinearity.

The two-level nature of the system is especially reflected in the second-order correlation function $g^{(2)}(t) = \langle a^\dagger(0)a^\dagger(t)a(t)a(0) \rangle / \langle a^\dagger(0)a(0) \rangle^2$, which describes the probability of detecting a second scattered photon at time t given a detection event at $t_0 = 0$. In steady-state, low pump regime $\Omega \ll \gamma$, it is straightforward to show that $g^{(2)}(0) = (1 + \gamma^{(2)}/\gamma)^{-2} = 7 \cdot 10^{-5}$. Thus, the disk exhibits an almost perfect anti-bunching effect characteristic of an ideal two-level system (to compare, $g^{(2)}(0) = 1$ in the absence of nonlinearities). The temporal duration of this anti-bunching dip around zero time delay $t = 0$ is given by $\sim 1/\gamma$ (in particular, within a two-level approximation $g^{(2)}(t) = 1 + e^{-\gamma t} - 2e^{-\gamma t/2}$ [31]), after which the system returns to a stationary value of $g^{(2)}(t) = 1$ as illustrated in Figure 3 (e). The sub-Poissonian nature of the scattered light, $g^{(2)}(0) < g^{(2)}(t)$, is a distinctly non-classical feature that reflects the inability of a two-level system to emit two excitations simultaneously [32].

While we have focused primarily on the absorptive nonlinearity, graphene will show also dispersive nonlinearities such as the Kerr effect [1]. Such a mechanism would in principle result in a shift in the resonance peak, in addition to the saturation of absorption in the graphene nano-ribbon. On the other hand, we do not expect that this would change our results in the case of a nano-disk, since this dispersive nonlinearity also leads to a plasmon blockade effect [21]. Interestingly, the estimated value of the dispersive interaction strength [21] is two orders of magnitude lower than the two-plasmon absorption rate.

In conclusion, we have shown that graphene possesses a strong nonlinear response in the form of multi-plasmon absorption, which leads to the saturation of absorption in graphene nano-ribbons, and the plasmon blockade effect in graphene nano-disks.

The authors would like to acknowledge the support of the European Commission and the Croatian Ministry of Science, Education and Sports Co-Financing Agreement No. 291823, Marie Curie FP7-PEOPLE-2011-COFUND NEWFELPRO project GRANQO, and FP7-ICT-2013-

* Electronic address: mjablan@phy.hr

- [1] R.W. Boyd, *Nonlinear Optics* (Academic Press, Amsterdam, 2008).
- [2] D.E. Chang, V. Vuletić, and M.D. Lukin, *Quantum nonlinear optics - photon by photon*, Nat. Photonics **8**, 685 (2014).
- [3] M. Kauranen and A.V. Zayats, *Nonlinear plasmonics*, Nat. Photonics **6**, 737 (2012).
- [4] K.S. Novoselov, A.K. Geim, S.V. Morozov, D. Jiang, Y. Zhang, S.V. Dubonos, I.V. Grigorieva, and A.A. Firsov, *Electric Field Effect in Atomically Thin Carbon Films*, Science **306**, 666 (2004).
- [5] B. Wunsch, T. Stauber, F. Sols, and F. Guinea, *Dynamical polarization of graphene at finite doping*, New. J. Phys. **8**, 318 (2006).
- [6] E.H. Hwang and S. Das Sarma, *Dielectric function, screening, and plasmons in two-dimensional graphene*, Phys. Rev. B **75**, 205418 (2007).
- [7] M. Jablan, H. Buljan, and M. Soljačić, *Plasmonics in graphene at infrared frequencies*, Phys. Rev. B **80**, 245435 (2009).
- [8] M. Jablan, M. Soljačić, and H. Buljan, *Plasmons in Graphene: Fundamental Properties and Potential Applications*, Proceedings of the IEEE **101**, 1689 (2013).
- [9] L. Ju, B. Geng, J. Horng, C. Girit, M. Martin, Z. Hao, H.A. Bechtel, X. Liang, A. Zettl, Y.R. Shen, and F. Wang, *Graphene plasmonics for tunable terahertz metamaterials*, Nat. Nanotechnology **6**, 630 (2011).
- [10] H. Yan, X. Li, B. Chandra, G. Tulevski, Y. Wu, M. Freitag, W. Zhu, P. Avouris, and F. Xia, *Tunable infrared plasmonic devices using graphene/insulator stacks*, Nature Nanotechnology **7**, 330 (2012).
- [11] J. Chen, M. Badioli, P. Alonso-Gonzalez, S. Thongrattanasiri, F. Huth, J. Osmond, M. Spasenović, A. Centeno, A. Pesquera, P. Godignon, A.Z. Elorza, N. Camara, F.J. Garcia de Abajo, R. Hillenbrand, and F.H.L. Koppens, *Optical nano-imaging of gate-tunable graphene plasmons*, Nature **487**, 77 (2012).
- [12] Z. Fei, A.S. Rodin, G.O. Andreev, W. Bao, A.S. McLeod, M. Wagner, L.M. Zhang, Z. Zhao, M. Thiemens, G. Dominguez, M.M. Fogler, A.H. Castro Neto, C.N. Lau, F. Keilmann, and D.N. Basov, *Gate-tuning of graphene plasmons revealed by infrared nano-imaging*, Nature **487**, 82 (2012).
- [13] S.A. Mikhailov and K. Zeigler, *Nonlinear electromagnetic response of graphene: frequency multiplication and the self-consistent-field effects*, J. Phys.: Condens. Matter **20**, 384204 (2008).
- [14] A.R. Wright, X.G. Xu, J.C. Cao, and C. Zhang, *Strong nonlinear optical response of graphene in the terahertz regime*, Appl. Phys. Lett. **95**, 072101 (2009).
- [15] K.L. Ishikawa, *Nonlinear optical response of graphene in time domain*, Phys. Rev. B **82**, 201402(R) (2010).
- [16] Z. Sun, T. Hasan, F. Torrisi, D. Popa, G. Privitera, F. Wang, F. Bonaccorso, D.M. Basko, and A.C. Ferrari, *Graphene Mode-Locked Ultrafast Laser*, ACS Nano **4**, 803 (2010).
- [17] E. Hendry, P.J. Hale, J. Moger, A.K. Savchenko, and S.A. Mikhailov, *Coherent Nonlinear Optical Response of Graphene*, Phys. Rev. Lett. **105**, 097401 (2010).
- [18] S.A. Mikhailov, *Theory of the giant plasmon-enhanced second-harmonic generation in graphene and semiconductor two-dimensional electron systems*, Phys. Rev. B **84**, 045432 (2011).
- [19] H. Yang, X. Feng, Q. Wang, H. Huang, W. Chen, A.T.S. Wee, and W. Ji, *Giant Two-Photon Absorption in Bilayer Graphene*, Nano Lett. **11**, 2622 (2011).
- [20] X. Yao and A. Belyanin, *Giant Optical Nonlinearity of Graphene in a Strong Magnetic Field*, Phys. Rev. Lett. **108**, 255503 (2012).
- [21] M. Gullans, D.E. Chang, F.H.L. Koppens, F.J. Garcia de Abajo, and M.D. Lukin, *Single-Photon Nonlinear Optics with Graphene Plasmons*, Phys. Rev. Lett. **111**, 247401 (2013).
- [22] M.T. Manzoni, I. Silvero, F.J. Garcia de Abajo, and D.E. Chang, *Second-order quantum nonlinear optical processes in graphene nanostructures*, arXiv:1406.4360v1 (2014).
- [23] J.D. Cox and F.J. Garcia de Abajo, *Electrically tunable nonlinear plasmonics in graphene nanoislands*, Nat. Communications **5**, 5725 (2014).
- [24] I. Al-Naib, J.E. Sipe, and M.M. Dignam, *High harmonic generation in undoped graphene: Interplay of inter- and intraband dynamics*, Phys. Rev. B **90**, 245423 (2014).
- [25] P.R. Wallace, *The Band Theory of Graphite*, Phys. Rev. **71**, 622 (1947).
- [26] See Supplemental Material.
- [27] H. Ehrenreich and M.H. Cohen, *Self-Consistent Field Approach to the Many-Electron Problem*, Phys. Rev. **115**, 786 (1959).
- [28] Z. Fang, S. Thongrattanasiri, A. Schlather, Z. Liu, L. Ma, Y. Wang, P.M. Ajayan, P. Nordlander, N.J. Halas, and F.J. Garcia de Abajo, *Gated Tunability and Hybridization of Localized Plasmons in Nanostructured Graphene*, ACS Nano **7**, 2388 (2013).
- [29] F.J. Garcia de Abajo, *Graphene Plasmonics: Challenges and Opportunities*, ACS Photonics **1**, 135 (2014).
- [30] S. Thongrattanasiri, F.H.L. Koppens, and F.J. Garcia de Abajo, *Complete Optical Absorption in Periodically Patterned Graphene*, Phys. Rev. Lett. **108**, 047401 (2012).
- [31] M.O. Scully and M.S. Zubairy, *Quantum optics* (University Press, Cambridge, 1997).
- [32] L. Mandel, *Non-Classical States of the Electromagnetic Field*, Phys. Scr. **T12**, 34 (1986).

Supplemental Material for "Multi-plasmon absorption in graphene"

The electron-plasmon interaction can be described by a Hamiltonian: $H_p = -\frac{1}{2}e\varphi_p(\omega, q)\hat{\rho}_q^\dagger e^{-i\omega t} + c.c.$, where φ_p is the plasmon potential at frequency ω and wavevector q , while $\hat{\rho}_q = \sum_j e^{-i\mathbf{q}\cdot\mathbf{r}_j}$ is the Fourier transform of the many-particle density operator $\hat{\rho}(\mathbf{r}) = \sum_j \delta(\mathbf{r} - \mathbf{r}_j)$ [S1]. To calculate N -plasmon absorption we can write the dissipated power as: $P^{(N)} = \sum_n N\hbar\omega \frac{dw_{n0}^{(N)}}{dt}$, and use the Fermi golden rule to find the probability of absorbing N plasmon quanta [S2]:

$$\frac{dw_{n0}^{(1)}}{dt} = \frac{2\pi}{\hbar} |\langle n|H_p|0\rangle|^2 \delta(E_n - E_0 - \hbar\omega) \quad (S1)$$

$$\frac{dw_{n0}^{(2)}}{dt} = \frac{2\pi}{\hbar} \left| \sum_m \frac{\langle n|H_p|m\rangle \langle m|H_p|0\rangle}{E_m - E_0 - \hbar\omega} \right|^2 \delta(E_n - E_0 - 2\hbar\omega) \quad (S2)$$

$$\frac{dw_{n0}^{(3)}}{dt} = \frac{2\pi}{\hbar} \left| \sum_{m,l} \frac{\langle n|H_p|m\rangle \langle m|H_p|l\rangle \langle l|H_p|0\rangle}{(E_m - E_0 - 2\hbar\omega)(E_l - E_0 - \hbar\omega)} \right|^2 \delta(E_n - E_0 - 3\hbar\omega) \quad (S3)$$

Here $|n\rangle$ is the many-body excited state of momentum $N\hbar\mathbf{q}$ and energy $N\hbar\omega$, with respect to the ground state $|0\rangle$, and we assume that the system is at zero temperature. In the most simple description of plasmons, we factorize our many-body state into the product of single-particle states which evolve in the mean-field, self-consistent plasmon potential $\varphi_p(\omega, q) = -\frac{e}{2\bar{\epsilon}_r\epsilon_0q} \langle \hat{\rho}_q \rangle$ [S1, S3, S4].

Low energy single-particle states in graphene are described by Dirac wavefunctions $\psi_{n\mathbf{k}}(\mathbf{r}) = \langle \mathbf{r}|n\mathbf{k} \rangle = \frac{1}{\sqrt{2A}} \begin{pmatrix} n \\ e^{i\mathbf{k}\cdot\mathbf{r}} \end{pmatrix}$ with linear dispersion $E_{n\mathbf{k}} = n\hbar v_F |\mathbf{k}|$, where \mathbf{k} is the electron wavevector, $n = \pm 1$, and we need to include the two-valley and two-spin degeneracy of the ground state [S5]. Finally, the damping rate can be written as $\gamma^{(N)} = P^{(N)}/W$ where $W = \frac{1}{2}\bar{\epsilon}_r\epsilon_0q|\varphi_p(\omega, q)|^2 A$ is the plasmon electrostatic energy, and it is straightforward to show that:

$$\frac{\gamma^{(N)}}{\omega} = F^{(N)}(\omega) \left| \frac{E_p}{E_e} \right|^{2N-2} \quad (S4)$$

where $|E_p| = q|\varphi_p|$ is the plasmon field, E_e is the intrinsic electric field given in the main text, and $F^{(N)}$ are dimensionless functions of plasmon frequency given by:

$$F^{(1)}(\omega) = \frac{e^2}{\bar{\epsilon}_r\epsilon_0q} \frac{4\pi}{A} \sum_{n_1 n_2 \mathbf{k}} |\langle n_2 \mathbf{k} + \mathbf{q} | e^{i\mathbf{q}\cdot\mathbf{r}} | n_1 \mathbf{k} \rangle|^2 \delta(E_{n_2 \mathbf{k} + \mathbf{q}} - E_{n_1 \mathbf{k}} - \hbar\omega) f_{n_1 \mathbf{k}} (1 - f_{n_2 \mathbf{k} + \mathbf{q}}) \quad (S5)$$

$$F^{(2)}(\omega) = \left(\frac{\omega}{v_F q} \right)^6 \frac{E_F^3}{k_F^2} \frac{\pi^2}{A} \sum_{n_1 n_3 \mathbf{k}} \left| \sum_{n_2} \frac{\langle n_3 \mathbf{k} + 2\mathbf{q} | e^{i\mathbf{q}\cdot\mathbf{r}} | n_2 \mathbf{k} + \mathbf{q} \rangle \langle n_2 \mathbf{k} + \mathbf{q} | e^{i\mathbf{q}\cdot\mathbf{r}} | n_1 \mathbf{k} \rangle}{E_{n_2 \mathbf{k} + \mathbf{q}} - E_{n_1 \mathbf{k}} - \hbar\omega} \right|^2 \times \delta(E_{n_3 \mathbf{k} + 2\mathbf{q}} - E_{n_1 \mathbf{k}} - 2\hbar\omega) f_{n_1 \mathbf{k}} (1 - f_{n_3 \mathbf{k} + 2\mathbf{q}}) \quad (S6)$$

$$F^{(3)}(\omega) = \left(\frac{\omega}{v_F q} \right)^{10} \frac{E_F^5}{k_F^2} \frac{3\pi^2}{2^5 A} \sum_{n_1 n_4 \mathbf{k}} \left| \sum_{n_2 n_3} \frac{\langle n_4 \mathbf{k} + 3\mathbf{q} | e^{i\mathbf{q}\cdot\mathbf{r}} | n_3 \mathbf{k} + 2\mathbf{q} \rangle \langle n_3 \mathbf{k} + 2\mathbf{q} | e^{i\mathbf{q}\cdot\mathbf{r}} | n_2 \mathbf{k} + \mathbf{q} \rangle \langle n_2 \mathbf{k} + \mathbf{q} | e^{i\mathbf{q}\cdot\mathbf{r}} | n_1 \mathbf{k} \rangle}{(E_{n_3 \mathbf{k} + 2\mathbf{q}} - E_{n_1 \mathbf{k}} - 2\hbar\omega)(E_{n_2 \mathbf{k} + \mathbf{q}} - E_{n_1 \mathbf{k}} - \hbar\omega)} \right|^2 \times \delta(E_{n_4 \mathbf{k} + 3\mathbf{q}} - E_{n_1 \mathbf{k}} - 3\hbar\omega) f_{n_1 \mathbf{k}} (1 - f_{n_4 \mathbf{k} + 3\mathbf{q}}) \quad (S7)$$

Here A is the surface area of the graphene flake, k_F is the Fermi wavevector, E_F is the Fermi energy, and $f_{n\mathbf{k}} = \Theta(E_F - E_{n\mathbf{k}})$ is the Fermi-Dirac distribution at zero temperature. Also we have explicitly written the expression (S5) to be evident that $F^{(1)}(\omega) = 2\text{Im}\varepsilon(\omega, q)$, where $\varepsilon(\omega, q)$ is the dielectric function calculated in the Random Phase Approximation [S3, S4], while we have simplified expressions (S6) and (S7) by using the plasmon dispersion relation $q = \frac{2\pi\varepsilon_0\varepsilon_F\hbar^2}{e^2E_F}\omega^2$.

Note that even at the room temperature $T = 300$ K and doping $n = 10^{12} \text{ cm}^{-2}$, $E_F/kT \approx 4.5$ and our calculations based on the zero-temperature approximation will still be qualitatively valid. To be more specific, at energy $\hbar\omega_0 = E_F$ where we were looking at two-plasmon absorption, we will also see a small contribution from the single-plasmon absorption. However, this effect will be suppressed by roughly $e^{-\Delta/kT} \approx e^{-2.3}$ where $\Delta \approx E_F/2$ is the energy gap required to enter the regime of single-plasmon absorption (see Figure 1 (a)).

From a technical stand point, one might also worry about severe heating of the graphene flakes at high laser powers. For example, in the case of a graphene nano-disk, the saturation intensity is $I_s = 50 \text{ kW/cm}^2$, with the corresponding absorption cross section on resonance: $\sigma_a(\omega_0) \approx 0.25A$ (see Figure 3 (c)), which would induce heating of the disk at $I_s\sigma_a/A = 13 \text{ kW/cm}^2$. However, note that graphene was reported to heat up at a rate $R = 3.3 \text{ K cm}^2/\text{kW}$ when subjected to a high direct current flow, while sitting on a room temperature SiO_2 substrate which acted as a heat sink [S6]. Thus, even in the continuous wave regime a laser would increase the temperature of the disk by only $\Delta T \approx 40$ K. Moreover in the case of ribbons, the saturation intensity, and the induced heating is even smaller.

* Electronic address: mjablan@phy.hr

- [S1] D. Pines, *The Theory of Quantum Liquids* (Benjamin, New York, 1966).
- [S2] J.J. Sakurai, *Modern Quantum Mechanics* (Addison-Wesley Publishing Company, Reading, Massachusetts, 1994).
- [S3] B. Wunsch, T. Stauber, F. Sols, and F. Guinea, *Dynamical polarization of graphene at finite doping*, New. J. Phys. **8**, 318 (2006).
- [S4] E.H. Hwang and S. Das Sarma, *Dielectric function, screening, and plasmons in two-dimensional graphene*, Phys. Rev. B **75**, 205418 (2007).
- [S5] P.R. Wallace, *The Band Theory of Graphite*, Phys. Rev. **71**, 622 (1947).
- [S6] M. Freitag, M. Steiner, Y. Martin, V. Perebeinos, Z. Chen, J.C. Tsang, and P. Avouris, *Energy Dissipation in Graphene Field-Effect Transistors*, Nano Lett. **9**, 1883 (2009).